DOCUMENT RESUME

ED 036 522 24 TE 001 728

AUTHOR	BLOUNT, NATHAN S.; AND OTHERS
TITLE	MEASURES OF WRITING MATURITY FROM TWO 500-WORD WRITING SAMPLES. REPORT FROM THE INDIVIDUALLY GUIDED INSTRUCTION IN ENGLISH LANGUAGE, COMPOSITION, AND LITERATURE PROJECT.
INSTITUTION	WISCONSIN UNIV., MADISON. RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING.
SPONS AGENCY	OFFICE OF EDUCATION (DHEW), WASHINGTON, D.C. BUREAU OF RESEARCH.
REPORT NO	TR-97
BUREAU NO	BR-5-0216
PUB DATE	SEP 69
CONTRACT	OEC-5-10-154
NOTE	14P.
EDRS PRICE	EDRS PRICE MF-$0.25 HC-$0.80
DESCRIPTORS	COMPOSITION (LITERARY), *COMPOSITION SKILLS (LITERARY), EVALUATION CRITERIA, *EVALUATION METHODS, GRADE 8, MATURITY TESTS, *MEASUREMENT TECHNIQUES, *RELIABILITY, *SAMPLING, SENTENCE STRUCTURE

ABSTRACT

THE PURPOSE OF THIS STUDY WAS TO ESTIMATE THE RELIABILITY OF VARIOUS MEASURES OF WRITING BEHAVIOR USING 500-WORD SAMPLES INSTEAD OF 1000-WORD SAMPLES. THE THEMES OF 135 EIGHTH GRADERS ON NARRATIVE AND EXPOSITORY TOPICS WERE COLLECTED FOR 6 WEEKS, ONE THEME PER WEEK. SAMPLES OF AT LEAST 1000 WORDS WERE OBTAINED FROM EACH STUDENT, AND THE CRITERIA OF HUNT (1965) WERE USED TO DETERMINE WHICH SENTENCES AND T-UNITS TO INCLUDE FOR TABULATION. TWO 500-WORD DIVISIONS OF THE 1000-WORD SAMPLES WERE MADE AND TWO COMPARISONS BETWEEN THE DIVISIONS WERE CARRIED OUT. THE FIRST COMPARISON INVOLVED SPLITTING THE SAMPLE AT THE SENTENCE OR T-UNIT NEAREST THE 500-WORD MIDPOINT. THE SECOND COMPARISON WAS BETWEEN THE ODD PAGES AND THE EVEN PAGES. ADJECTIVE, ADVERB, AND NOUN CLAUSES OCCURRED FREQUENTLY ENOUGH TO BE OF IMPORTANCE AND CORRELATED .35, .31, AND .08, RESPECTIVELY, IN THE FIRST-HALF/SECOND-HALF SAMPLES, AND .46, .42, AND .26 IN THE ODD/EVEN PAGE SAMPLES. THE FIVE SYNOPSIS SCORES (CLAUSE LENGTH, CLAUSES PER T-UNIT, T-UNIT LENGTH, T-UNITS PER SENTENCE, AND SENTENCE LENGTH) CORRELATED .48, .22, .56, .48, AND .62 BETWEEN FIRST-HALF AND SECOND-HALF RESPECTIVELY, AND .69, .54, .74, .65, AND .77 BETWEEN ODD-PAGE/EVEN-PAGE SAMPLES. (THREE TABLES ARE INCLUDED.) (AUTHOR/LH)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR-5-0216-TR97
PA24
OE/BR

ED036522

Technical Report No. 97

MEASURES OF WRITING MATURITY FROM TWO 500-WORD WRITING SAMPLES

by
Nathan S. Blount, Wayne C. Fredrick, & Shelby L. Johnson

Report from the Individually Guided Instruction in English Language, Composition, and Literature Project

Nathan S. Blount and Lester S. Golub, Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

September 1969

TE 001 728

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Henry Chauncey
President
Educational Testing Service

Elizabeth Koontz
President
National Education Association

Patrick Suppes
Professor
Department of Mathematics
Stanford University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Roderick McPhee
President
Punahou School, Honolulu

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

Francis S. Chase
Professor
Department of Education
University of Chicago

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

G. Wesley Sowards
Director, Elementary Education
Florida State University

UNIVERSITY POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

John Guy Fowlkes
Director
Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

M. Crawford Young
Associate Dean
The Graduate School

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Donald J. McCarty
Dean
School of Education

***James W. Cleary**
Vice Chancellor for Academic Affairs

H. Clifton Hutchins
Chairman
Department of Curriculum and Instruction

Ira Sharkansky
Associate Professor of Political Science

Leon D. Epstein
Dean
College of Letters and Science

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Max R. Goodson
Professor of Educational Policy Studies

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Ronald R. Allen
Associate Professor of Speech and of Curriculum and Instruction

Gary A. Davis
Associate Professor of Educational Psychology

Max R. Goodson
Professor of Educational Policy Studies

Richard G. Morrow
Assistant Professor of Educational Administration

Vernon L. Allen
Associate Professor of Psychology
(On leave 1968-69)

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Warren O. Hagstrom
Professor of Sociology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Nathan S. Blount
Associate Professor of English and of Curriculum and Instruction

Frank H. Farley
Assistant Professor of Educational Psychology

John G. Harvey
Associate Professor of Mathematics and Curriculum and Instruction

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Robert C. Calfee
Associate Professor of Psychology

John Guy Fowlkes (Advisor)
Professor of Educational Administration
Director of the Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Thomas A. Romberg
Assistant Professor of Mathematics and of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

Burton W. Kreitlow
Professor of Educational Policy Studies and of Agricultural and Extension Education

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

MANAGEMENT COUNCIL

***Herbert J. Klausmeier**
Director, R & D Center
Acting Director, Program 1

Thomas A. Romberg
Director
Programs 2 and 3

James E. Walter
Director
Dissemination Section

Dan G. Woolpert
Director
Operations and Business

Mary R. Quilling
Director
Technical Section

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Individually Guided Instruction in English Language, Composition, and Literature Project in Program 2. General objectives of the Program are to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to develop assessment procedures for those concepts and skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to these Program objectives, the long-range objective of the English Project is to install and test materials for individually guided instruction in language, composition, and literature. Prerequisite activities include formulating behavioral objectives for students and teachers, based on a content and concepts outline, and developing measurement instruments related to the behavioral objectives.

CONTENTS

		Page
	List of Tables	vii
	Abstract	ix
I	The Problem	1
II	Experimental Procedure	2
III	Results and Discussion	4
	References	9

LIST OF TABLES

Table		Page
1	Means and Standard Deviations Tabulated from the 500-Word Writing Samples of 135 Eighth Graders	4
2	Correlations for 17 Pairs of Writing Measures from Two 500-Word Samples	5
3	Intercorrelations of the Five Synopsis Scores in the Odd-Even Comparison	6

ABSTRACT

The purpose of this study was to estimate the reliability of various measures of writing behavior using 500-word samples instead of 1000 words. The themes of 135 Eighth Graders were collected over a period of 6 weeks, one theme per week. The topics, expository and narrative, were selected jointly by teachers and students as part of the normal curriculum in English. Samples of at least 1000 words were obtained from each student and the criteria of Hunt (1965) were used to determine which sentences and T-units to include for tabulation. Two 500-word divisions of the 1000-word samples were made. The themes written during the first 3 weeks were compared to those of the second 3 weeks, and the odd pages were compared to the even pages.

Eight of the tabulated structures (clauses of comparison, deferred subject, cleft sentence, "the more, the merrier" construction, subjunctive word order, special which clauses, adjective complements, and adverb complements) occurred very infrequently and were essentially uncorrelated ($r < .26$) from sample to sample. Adjective, adverb, and noun clauses were more frequent and correlated .35, .31, and .08, respectively, in the first-half/second-half samples, and .46, .42, and .26 in the odd/even pages samples. The five synopsis scores (clause length, clauses per T-unit, T-unit length, T-units per sentence, and sentence length) correlated .48, .22, .56, .48, and .62 between first half and second half respectively, and .69, .54, .74, .65, and .77 between odd-page and even-page samples.

I
THE PROBLEM

Hunt (1965) used 1000-word writing samples from which he tabulated eleven syntactical structures and five synopsis scores. The work required to tabulate such a large amount of writing is arduous, and if a smaller sample can be used which will still give an accurate measure of writing behavior, much menial labor can be spared. The purpose of the present study was to provide some data on the reliability of the Hunt measures when 500 words of 1000-word samples are used.

Some very old studies have made inferences about what length a writing sample should be. Frogner (1933) thought that for some measures one needed to analyze 50,000 to 60,000 words before an accurate measure of group performance was achieved. She made no estimate of the sample size required from an individual. This high number of words was especially needed when tabulating errors and specific categories of clauses. Other measures, such as the proportions of simple, compound, and complex sentences, produced by groups seemed to stabilize after 10,000 words had been tabulated. Frogner also noted that the kind of writing influenced the complexity of the sentences. She showed that expository writing produced a higher percentage of complex sentences than narrative, which in turn produced a higher proportion of complex sentences than those produced in letter-writing.

Anderson (1937) tested the feasibility of using samples from individuals as small as 150 words to tabulate sentence length. He concluded that 150-word samples were too small to reliably measure even the most objective indices. He postulated that sentence length would vary with the situation and the subject matter.

Chotlos (1944) made a systematic comparison of the reliability of type-token ratios (number of unique words divided by the number of words) obtained from 100-, 500-, 1000-, and 1500-word samples written by individuals. The type-token ratios for two 100-word samples correlated .38, while two 500-word samples correlated .67, 1000-word .81, and 1500-word .83. Though these correlations showed a clear increase with sample size, there was some confounding involved since, as sample size increased, the number of words between samples decreased. For example, the 1500-word samples were the first and second half of a 3000-word corpus, but the 1000 words were the first and last thirds of this same corpus. Thus, 1000 words of writing had taken place between the two 1000-word samples, but none had occurred between the 1500-word samples. Whatever effect writing has on writing, it became greater with the smaller samples, until with 100 words a massive 2800 words had intervened between the two samples. The type-token ratio is also critically dependent upon the number of words in a way that other measures (sentence length, clause length, T-unit length) are not. The types of words are accumulated rapidly as one begins tabulating. As tabulating continues the rate at which new types of words are added declines, but word tokens are added at a constant rate throughout.

Recent studies have settled upon 1000 words per S as a suitable body of writing for purposes of analysis (Hunt, 1965; Blount, Fredrick, & Johnson, 1968). One can ask, however, whether a smaller sample will result in measures that approach the reliability of the same measures obtained from a larger sample. If a 500-word sample can replace a 1000-word sample with only a slight decrease in reliability, the savings of time and effort are obvious.

Taking under consideration the usual level of variability in measures of writing, it was decided that whenever a variable accounted for one-fourth of the variance in two successive samples ($r > .5$) that variable was of some value as a measure of group performance. Whenever one-half of the variance was identifiable ($r > .7$) the measure was of value for assessing individual performance.

II
EXPERIMENTAL PROCEDURE

The Ss were Eighth Graders in a large public junior high school in Wisconsin. High and low ability Ss were drawn from the Eighth Grade population (242 students in the classrooms of the participating teachers). Among these were 72 females and 63 males. Seventy-two Ss had above average scores (119-145) on the California Test of Mental Maturity and 63 were below average (90-106).

During the first part of the 1966-67 school year each S wrote one in-class theme per week as part of the normal work in his English class. Ss did not know that the themes were being collected from the teachers, copied, and returned by the experimenters. The topics of the themes were the usual descriptive, narrative, and expository assignments decided upon jointly by the students and teachers. The successive weekly themes were collected until each of the 135 Ss had produced at least 1000 words of text. Generally this was accomplished by about the sixth theme.

From the copies the themes were typed triple-spaced in preparation for the tabulation of various measures of writing ability. The guidelines developed by Hunt (1965) were used in tabulating 17 measures, which included the following eleven subordinate structures, the total amount of subordination, and five synopsis scores:

1. Adjective clauses: clause structures which functioned as adjectives in modifying a noun or pronoun.
   <That was the place where he lost his keys.>
2. Adverb clauses: clause structures which functioned as movable adverbs.
   <I heard noises when I woke up.>
3. Noun clauses: clause structures which occurred in the nominal positions of subject, direct object, object of preposition, and appositive.
   <That there are good teenagers is never mentioned.>
4. Clauses of comparison: adverb-like clauses which were not usually movable and which used comparatives such as so that, more than, as long as, etc.
   <The clouds seemed so low that you could catch one.>
5. Deferred subject: clause structures which were the "logical" subjects of the sentence but which followed the "grammatical" subject it.
   <It was good that such a thing happened.>
6. Cleft sentence: clause structures (other than "logical" subjects) caused by it as the "grammatical" subject.
   <It's the fog that keeps the fumes from escaping.>
7. "The more, the merrier": sentences which involved a the more this, the more that statement.
   <The more he worried, the more he felt sorry for himself.>
8. Subjunctive: use of different word order to express the subjunctive mood.
   <I would have been killed had it not been for him.>
9. Special which clauses: which clauses that modified an entire sentence.
   <When we come in the morning we can go to the gym, which I think is nice.>
10. Adjective complement: clauses that complemented an adjective and could not be moved.
    <She was sure that I had left home.>
11. Adverb complement: clauses that followed verbs like seem, appeared, feel, etc.
    <She looked as if she had seen a ghost.>
12. All subordinate clauses: the total of 1 - 11 above.

For the synopsis scores, the number of words, clauses, T-units, and sentences were tabulated according to these guidelines: Words were all letter sequences that, regardless of spelling, could be considered a unit of speech. Contractions were counted as two words, and proper nouns (Yellowstone National Park) were counted as one word. Clauses were all main and subordinate structures that contained a subject and a finite verb phrase. T-units were defined as "one main clause plus the subordinate clauses attached to or embedded within it [Hunt, 1965, p. 49]." Sentences were the words written between a capital letter and some end punctuation. As an example of each of these terms, note the following sentence which contains three T-units marked by double slashes, a main clause within each T-unit, three subordinate clauses marked by a single slash and a symbol, and 42 words: < If I had a million dollars (adv) / I think / (N) that I would do a lot of things, // but I would put most of the money in the First National Bank / / and I would start a job / (adv) so people wouldn't call me a lazy man. //> The tabulations of words, clauses, T-units, and sentences were then used to compute these synopsis scores:

13. Clause length: words per clause.

14. Clauses per T-unit.

15. T-unit length: words per T-unit.

16. T-units per sentence.

17. Sentence length: words per sentence.

Under Hunt's guidelines all sentence fragments, garbles, direct discourse, and questions were deleted from consideration. The 17 measures were tabulated independently by two raters who had been trained previously and had used the tabulating system in previous work for the experimenters. In past work the correlations between the two raters had ranged from .84 to 1.00 on the various indices tabulated. All discrepancies between the two raters were resolved by the experimenter (SJ) who had originally trained the raters. A detailed discussion of each measure can be found in Blount, Fredrick, and Johnson (1968).

When typed sentence-by-sentence and triple-spaced, each S's writing sample became a document of about 12 pages. Two comparisons between halves of the 1000-word samples were then made. The first comparison involved splitting the sample at the sentence or T-unit nearest the 500-word midpoint. Each of the 17 scores obtained from the first half was correlated with the respective score from the second half. The second comparison was between the odd pages of the typed sample and the even pages. Each of the 17 scores from the odd pages (to the end of the sentence or T-unit nearest 500 words) was correlated with the respective score from the even pages.

III
RESULTS AND DISCUSSION

The means and standard deviations from 135 odd-page samples from the odd-even pages split are presented in Table 1. The equivalent tables for even-page samples and first-half/second-half samples are not presented since all four tables were practically identical. The means presented include eleven tabulations of subordinate structures, the total of all eleven kinds of subordinate structures, and five synopsis scores. In 500 words of Eighth-Grade writing, an average of 22 subordinate structures appeared; roughly 5 of these were adjective clauses, 9 adverb clauses, and 7 noun clauses. Each clause averaged about eight words in length, and appeared with a frequency of three clauses per every two T-units. T-units themselves averaged 13 words long, and one-third of the time there was more than one T-unit in a sentence. Sentence length averaged more than 16 words. Each of these statistics closely approximated the results found in previous studies (Hunt, 1965; Blount, et al., 1968). [Adjustments were made to frequency counts so that 500- and 1000-word samples could be compared.]

Table 1

Means and Standard Deviations Tabulated from the 500-Word Writing Samples of 135 Eighth Graders

Variable	All 135 Ss		63 Males		72 Females		72 High IQ		63 Low IQ	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	Mean	sd
Adjective Clauses	5.3	2.9	5.4	3.1	5.2	2.7	5.5	3.0	5.1	2.8
Adverb Clauses	8.8	4.2	8.3	4.2	9.3	4.2	8.4	4.2	9.3	4.3
Noun Clauses	7.0	3.8	6.7	4.0	7.2	3.6	6.8	3.8	7.2	3.8
Clauses of Comparison	.4	.7	.5	.7	.3	.7	.4	.7	.4	.7
Deferred Subject	.1	.3	.0	.2	.1	.4	.1	.3	.1	.4
Cleft Sentence	.1	.3	.1	.4	.1	.3	.1	.3	.1	.4
"The More, the Merrier"	.0	.1	.0	.2	.0	-	.0	.2	.0	-
Subjunctive	.0	.2	.0	.3	.0	-	.0	-	.0	.3
Special Which Clauses	.1	.2	.1	.2	.0	.2	.1	.2	.0	.2
Adjective Complement	.1	.4	.1	.3	.2	.5	.1	.4	.1	.5
Adverb Complement	.1	.2	.0	.2	.1	.3	.1	.2	.0	.2
All Subordinate Clauses	22.0	6.2	21.3	6.6	22.6	5.9	21.6	6.6	22.5	5.7
Clause Length	8.2	1.2	8.4	1.3	8.1	1.1	8.3	1.3	8.2	1.1
Clauses per T-unit	1.6	.2	1.6	.2	1.6	.2	1.6	.2	1.6	.2
T-unit Length	13.1	2.7	13.2	2.7	13.0	2.7	13.1	2.9	13.1	2.5
T-units per Sentence	1.3	.2	1.3	.2	1.2	.2	1.2	.2	1.3	.2
Sentence Length	16.4	3.5	16.9	3.7	16.0	3.3	16.0	3.1	16.9	3.9

To compare the differences between odd/even page samples, t tests were computed on all 17 dependent measures. These tests were not significant nor were there significant differences in comparisons of the first-half/second-half samples. There were also no significant differences when the mean scores of the 63 males were compared to the means of the 72 females and neither were any significant differences observed in the comparison of high and low IQ Ss.

The feasibility of using 500-word samples to obtain measures of subordination and writing maturity can be seen more clearly in the correlation coefficients presented in Table 2. The Pearson product-moment correlations between the samples in the odd/even split for the 11 subordinate structures were far below acceptable levels for individual and group measurement. The first-half/second-half correlations for these 11 measures were even smaller than the corresponding correlations from the odd/even split. Thus, a 500-word sample did not give a reliable measure of the use of certain subordinate structures. Even when considering all subordinate structures, the correlation between two 500-word samples only reached .46. By using the Spearman-Brown prophecy formula, it was estimated that two 1000-word samples might correlate about .63 for total use of subordination, but below .59 for any measure of a particular subordinate structure. The 1000-word sample may, then, also be inadequate and too unreliable for measuring specific subordinate structures.

For the synopsis scores, however, the correlations between samples were higher, and for the odd/even pages split they approached a useful level. The measures of clause length, T-unit length, and sentence length appeared to have reached a level of reliability ($r > .69$) that would be sufficient to obtain reliable measures of a group's characteristic writing behavior. For individuals, one might still desire the slightly higher reliability than might be expected from 1000-word samples.

Table 2

Correlations for 17 Pairs of Writing Measures
Taken from Two 500-Word Samples

Syntactical Structures	Frequency in 1000 Words	Odd/Even Pages All Ss	First-Half/ Second-Half All Ss	Odd/Even Pages by Groups			
				Male Ss	Female Ss	High Ability	Low Ability
Adjective Clauses	10.49	.42	.35	.53	.31	.34	.52
Adverb Clauses	17.73	.41	.31	.46	.36	.52	.27
Noun Clauses	14.21	.26	.08	.24	.27	.45	.02
Clauses of Comparison	.93	-.01	-.12	.05	-.07	.01	-.03
Deferred Subject	.16	.01	.03	.25	-.06	-.08	.20
Cleft Sentence	.20	.29	.11	.46	.08	.23	.35
"The More, the Merrier"	.03	-.02	-.01	-.03	--	--	--
Subjunctive	.01	--	-.01	--	--	--	--
Special Which Clauses	.19	.08	-.05	-.09	.30	-.09	.22
Adjective Complement	.24	-.04	.10	-.05	-.07	-.04	-.07
Adverb Complement	.13	.04	.07	.19	-.06	.13	-.06
All Subordinate Clauses	44.33	.46	.39	.49	.43	.51	.39
Synopsis Scores	Mean						
Clause Length	8.32	.69	.48	.72	.64	.72	.64
Clauses per T-unit	1.59	.54	.22	.56	.52	.65	.42
T-unit Length	13.22	.74	.56	.81	.68	.82	.61
T-units per Sentence	1.25	.65	.48	.72	.51	.69	.62
Sentence Length	16.47	.77	.62	.76	.79	.79	.75

In comparing the two methods of sampling from a writing sample, odd/even pages vs. first-half/second-half, it is obvious that the odd/even pages procedure was superior. For the synopsis scores the odd/even correlations ranged from .54 to .77 compared to a range of .22 to .62 for the first-half/second-half. That is, sampling every other page represents the entire corpus of writing more adequately than if one considers only the first half. Perhaps the ideal procedure would be to consider every other sentence, or every third sentence, or every fourth sentence, etc., depending on the proportion of the total corpus one wanted to sample.

In Table 2 the correlations between writing samples taken from the odd/even split are presented for male, female, high IQ, and low IQ Ss separately. Several differences between pairs of correlations are apparent, but what these differences indicate about writing performance is uncertain. Possibly the writing of the male and high IQ Ss was more stable and consistent since the correlations for these two groups were generally higher than for female and low IQ Ss. Consistency does not, however, imply maturity.

In Table 3 the intercorrelations between the odd/even pages synopsis scores are presented. All correlations greater than .26 in absolute value show significant relationship ($p < .01$) for 133 degrees of freedom. Note that measures of clause length, clauses per T-unit, and T-units per sentence are largely independent measures of writing behavior. That is, a person who writes long clauses does not necessarily write more or less clauses than someone who writes short clauses. If clauses per T-unit is considered a measure of subordination, and T-units per sentence a measure of coordination, it is apparent from Table 3 that at Eighth Grade these two skills are uncorrelated. Clause length, sentence length, and T-unit length, however, are significantly correlated since sentences are made up of T-units and clauses, and T-units include any clauses present.

The following conclusions become apparent from the comparison of two 500-word samples:

1. Several structures (clauses of comparison, deferred subject, cleft sentence, "the more, the merrier," subjunctive, special which clauses, adjective complement, and adverb complement) were very infrequent in 500 words and also 1000 words. Much larger samples of writing covering various writing situations would be needed to make reliable estimates of their incidence.

2. Adjective, adverb, and noun clauses together occurred about 5, 9, and 7 times, respectively, in an average 500-word sample of Eighth-Grade writing. The frequencies correlated from .26 to .42 from sample to sample (odd/even pages). These correlations seem too low for most purposes, and even the

Table 3

Intercorrelations of the Five Synopsis Scores in the Odd/Even Comparison*

Variable		Odd Pages W/C	C/T	W/T	T/S	W/S	Even Pages W/C	C/T	W/T	T/S	W/S
Odd Pages											
Clause Length	(W/C)	1.00	-.03	.77	-.22	.53	.69	.12	.58	-.28	.32
Clauses per T-unit	(C/T)		1.00	.60	-.05	.51	.09	.54	.45	.03	.44
T-unit Length	(W/T)			1.00	-.21	.74	.59	.45	.74	-.20	.54
T-units per Sentence	(T/S)				1.00	.49	-.01	-.06	-.04	.65	.42
Sentence Length	(W/S)					1.00	.53	.34	.63	.27	.77
Even Pages											
Clause Length	(W/C)						1.00	-.05	.67	-.15	.52
Clauses per T-unit	(C/T)							1.00	.70	-.24	.47
T-unit Length	(W/T)								1.00	-.28	.72
T-units per Sentence	(T/S)									1.00	.45
Sentence Length	(W/S)										1.00

* $r \geq .26$ is significant at the .01 level

higher estimates from a larger sample (1000-words might correlate .41 to .59) would be inadequate for measures of individual performance.

3. The total number of all subordinate structures reached a sample-to-sample reliability of .46, also too unreliable for most purposes.

4. The synopsis scores showed relatively high levels of reliability compared to the subordinate structure frequencies.

5. Though groups of Ss did not differ significantly from sample to sample, different sampling procedures varied in how well they represented a larger sample. In the present case, comparing odd to even pages resulted in higher correlations than comparing first-half to second-half samples.

Thus, it appears that at least three of the synopsis scores (clause length, clauses per T-unit, and T-units per sentence) are non-redundant measures that may be obtained objectively and reliably from a writing sample of 500 words. Two other synopsis scores, sentence length and T-unit length, are also reliable when obtained from 500-word samples, but reflect factors which are already measured by the other three synopsis scores. Sentence length and T-unit length are therefore correlated more highly with the other synopsis scores.

The reliability of the synopsis scores from 500-word samples appears to compare favorably to the levels usually obtained in the measurement of writing. Apparently, sampling systematically (every other page or sentence) will provide a reliable representation of the larger corpus of writing, as far as the synopsis scores are concerned. For measures of the frequency of specific types of clauses, the only option appears to be the tabulation of very large writing samples over many types of topics and situations.

GPO 814-474-3

REFERENCES

Anderson, J. E. An evaluation of various indices of linguistic development. Child Development, 1937, 8, 62-68.

Blount, N. S., Fredrick, W. C., & Johnson, S. L. The effect of a study of grammar on the writing of eighth-grade students. Technical Report from the Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1968, No. 69.

Chotlos, J. W. Studies in language behavior. IV. A statistical and comparative analysis of individual written language samples. Psychological Review (Psychological Monographs), 1944, 56, 77-111.

Frogner, E. Problems of sentence structure in pupils' themes. English Journal, 1933, 22, 742-749.

Hunt, K. W. Grammatical structures written at three grade levels. National Council of Teachers of English Research Report No. 3. Champaign, Illinois: National Council of Teachers of English, 1965.

GPO 814-474-2